



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/608,591	06/27/2003	James M. Sweet	D/A2555Q1	8445

25453 7590 08/10/2005

PATENT DOCUMENTATION CENTER  
XEROX CORPORATION  
100 CLINTON AVE., SOUTH, XEROX SQUARE, 20TH FLOOR  
ROCHESTER, NY 14644

EXAMINER
----------

HILLERY, NATHAN

ART UNIT	PAPER NUMBER
----------	--------------

2176

DATE MAILED: 08/10/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

# Office Action Summary

Application No.

10/608,591

Applicant(s)

SWEET ET AL.

Examiner

Nathan Hillery

Art Unit

2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☒ Responsive to communication(s) filed on 06 October 2003.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1-15 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-15 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 27 June 2003 is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)                        | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)               | Paper No(s)/Mail Date. _____  |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| Paper No(s)/Mail Date <u>6/27/03</u> .   | 6) <input type="checkbox"/> Other: _____                                    |

### **DETAILED ACTION**

1. This action is responsive to communications: Oath or Declaration filed on 10/6/03.
2. Claims 1 – 15 are pending in the case. Claims 1, 6 and 11 are independent.

### ***Drawings***

3. New corrected drawings in compliance with 37 CFR 1.121(d) are required in this application because there are two sets of drawings submitted on the same day – one with figs 1 – 6 and another with Figs 1 – 8; consequently, the office does not know which set is correct or should be considered for examination. Applicant is advised to employ the services of a competent patent draftsman outside the Office, as the U.S. Patent and Trademark Office no longer prepares new drawings. The corrected drawings are required in reply to the Office action to avoid abandonment of the application. The requirement for corrected drawings will not be held in abeyance.

### **INFORMATION ON HOW TO EFFECT DRAWING CHANGES**

#### **Replacement Drawing Sheets**

Drawing changes must be made by presenting replacement sheets which incorporate the desired changes and which comply with 37 CFR 1.84. An explanation of the changes made must be presented either in the drawing amendments section, or remarks, section of the amendment paper. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). A replacement sheet must include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of the amended drawing(s) must not be labeled as "amended." If the changes to the drawing figure(s) are not accepted by the examiner, applicant will be notified of any required corrective action in the next Office action. No further drawing submission will be required, unless applicant is notified.

Art Unit: 2176

Identifying indicia, if provided, should include the title of the invention, inventor's name, and application number, or docket number (if any) if an application number has not been assigned to the application. If this information is provided, it must be placed on the front of each sheet and within the top margin.

### **Annotated Drawing Sheets**

A marked-up copy of any amended drawing figure, including annotations indicating the changes made, may be submitted or required by the examiner. The annotated drawing sheet(s) must be clearly labeled as "Annotated Sheet" and must be presented in the amendment or remarks section that explains the change(s) to the drawings.

### **Timing of Corrections**

Applicant is required to submit acceptable corrected drawings within the time period set in the Office action. See 37 CFR 1.85(a). Failure to take corrective action within the set period will result in ABANDONMENT of the application.

If corrected drawings are required in a Notice of Allowability (PTOL-37), the new drawings **MUST** be filed within the **THREE MONTH** shortened statutory period set for reply in the "Notice of Allowability." Extensions of time may **NOT** be obtained under the provisions of 37 CFR 1.136 for filing the corrected drawings after the mailing of a Notice of Allowability.

### ***Claim Rejections - 35 USC § 103***

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. Claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 41, and 46 – 48 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and in further view of Raghavan et al. (US 6886129 B1).

6. **Regarding independent claim 11**, Bharat et al. teach that *we locate pages that point to at least one of the pages in the start set 201. We call this set of pages the back*

set 202. With the AltaVista search engine, "link:URL" queries can be used to identify back set pages for each start set page. We add one node 212 to the n-graph 211 for each page of the back set 202. Similarly, the pages pointed to by the start set 201 are located. This can be done by fetching each start set page and extracting the hyperlinks in each of the pages. The pages pointed to by the hyperlinks constitute the forward set 203. Nodes for the forward set of pages are also added to the n-graph 211. Thus, the input set of pages 204 includes the back, start, and forward sets 201-203. The input set 204 includes pages which do not directly satisfy the query, i.e., pages that do not include key words exactly as specified in the query. However, these pages may be useful because they are linked to pages of the start set. A larger n-graph 211 can be constructed by repeating this process for the back and forward sets 202-203 to add more indirectly linked pages. At this stage, the n-graph 211 has nodes 212 but no edges. After we have constructed the nodes 212, we add the directed edges 213. If a link points to a page that is represented by a node in the graph, and both pages are on different servers, then a corresponding edge 213 is added to the graph 211. Nodes representing pages on the same server are not linked. This prevents a single Web site with many self-referencing pages to unduly influence the outcome. This completes the n-graph 211 (Column 4, line 61 – Column 5, line 20), compare with **performing a page-level link analysis that identifies those hyperlinks on a page linking to a candidate document page further comprising a methodology of: searching page data to create a list of links in the document; analyzing each link in conjunction with each other link in the list of links to identify link pairings; assembling link pairings in**

**order to form clusters of links; and, examining the links in the cluster of links for locality; and, performing a recursive application of the page-level link analysis to the linked candidate document page and any further nested candidate document pages thereby identified, until a collective set of identified candidate document pages is assembled.** Bharat et al. do not explicitly teach **performing a document-level analysis that examines the collective set of identified candidate document pages for grouping into one or more documents.** However, Raghavan et al. teach that *the present invention provides a method for identifying and enumerating groups of pages of common interest from a collection of hyper-linked pages, including the steps of: (a) identifying community cores from the collection where each core includes first and second sets of pages and each page in the first set points to every page in the second set; and (b) expanding each identified core into a full community which is a subset of the pages regarding a particular topic. To minimize the number of duplicate pages, in the hyper-links between any two pages on the same site are removed. In addition, the pages of more established sites are discarded because they might skew the results. Highly similar pages are replaced with a single page that is representative of the replaced pages, with the hyper-links previously pointing to the replaced pages now pointing to the representative page* (Column 4, lines 6 – 20), compare with **performing a document-level analysis that examines the collective set of identified candidate document pages for grouping into one or more documents.** It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the invention of Bharat et al. with that of Raghavan et al. because such a

combination would provide the users of Bharat et al. with *a method for identifying implicitly defined communities from a collection of hyper-linked pages* (Column 3, lines 61 – 63).

7. **Regarding dependent claim 12**, Bharat et al. teach that *we assign a similarity weight to each node 213 of the sub-graph 255. Various document similarity measuring techniques have been developed in Information Retrieval to determine the goodness of fit between a "target" document and a collection of documents. These techniques typically measure a similarity score based on word frequencies in the collection and a target document* (Column 6, lines 51 – 57), compare with **the step for analyzing each link further comprises determining a score for each link pairing, and the scoring is determined by a similarity criteria.**

8. **Regarding claims 1, 2, and 4**, the claims incorporate substantially similar subject matter as claims 11 – 15 and are rejected along the same rationale.

9. **Regarding claims 6, 7, and 9**, the claims incorporate substantially similar subject matter as claims 11 – 15 and are rejected along the same rationale.

10. Claims 3, 5, 8, 10, 13, and 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Raghavan et al. (US 6886129 B1) as applied to claims above, and in further view of Min et al. (US 6633868 B1).

11. **Regarding dependent claims 13 and 15**, neither Bharat et al. nor Raghavan et al. explicitly teach **the scoring is determined by a proximity criteria, and the scoring is determined by a regularity criteria.** Min et al. teach that *For each document,  $i$ , a*

*matrix  $D(i)$  is calculated in the same manner as the C-db, C. That is, the elements of the square matrix are determined by the proximity and frequency of word pairs. (C is in fact the summation of all  $D(i)$  matrices; thus both the C and  $D(i)$  may be pre-computed to decrease retrieval times.) The elements on the diagonal may be set to zero so that only relationships between word pairs are taken into account. Normalization factors may also be applied to adjust for parameters such as document length, word pair frequency, etc. The matrix product of  $D(i)$  and S (that is, an element-by-element multiplication, followed by a summation of all elements to produce a scalar) computes a weight,  $W(i)$ , that correlates with the number and proximity of relevant word pairs found in each document. The product should be limited to elements common to both matrices in order to minimize the total number of computations. The final document ranking, R, is a simple sorting of the weights from highest to lowest values:  $R = \text{Sort}[W]$  (Column 7, lines 38 – 55), compare with **the scoring is determined by a proximity criteria**, and **the scoring is determined by a regularity criteria**. It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the combined invention of Bharat et al. and Raghavan et al. with that of Min et al. because such a combination would allow the users of Min et al. the benefit of a *computer-implemented method for improving query-based document retrieval using the vast amount of contextual information (i.e., information about the relationships between words) within the document collection to be searched* (Column 2, lines 61 – 65).*

12. **Regarding claims 3, 5, 8, and 10**, the claims incorporate substantially similar subject matter as claims 13 and 15 and are rejected along the same rationale.



**Conclusion**

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Nathan Hillery whose telephone number is (571) 272-4091. The examiner can normally be reached on M - F, 10:30 a.m. - 7:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on (571) 272-4136. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

NH

*William L. Bashore*  
**WILLIAM BASHORE**  
**PRIMARY EXAMINER**  
8/7/2005